

## Nagyszótáras híryanagok felismerési pontosságának növelése morfémaalapú, folyamatos beszédfelismerővel

Tarján Balázs, Mihajlik Péter, Tüske Zoltán

Budapesti Műszaki és Gazdaságtudományi Egyetem,  
Távközlési és Médiainformatikai Tanszék  
{tarjanb, mihajlik, tuske}@tmit.bme.hu

**Kivonat:** Morfémaalapú beszédfelismerőnek tekintjük azokat a felismerőket, melyek szónál kisebb, morfémaszerű elemekre épülő nyelvi modellt használnak. Kísérleteink során öt különböző szegmentáló eljárással készített morfémaalapú felismerő teljesítményét hasonlítottuk egy standard, szó alapú rendszeréhez tervezett beszédű, híryanag felolvasásos feladaton. Megállapítottuk, hogy mind statisztikai, mind szabályalapú szegmentáló algoritmust használva, morféma alapon jelentős mértékben növelni lehet a felismerési pontosságot. Különösen alacsony hibaarányt értünk el egy hibrid eljárással, mely a statisztikai módszert nyelvspecifikus tudással egészíti ki. Felügyelet nélküli beszélőadaptációs technológiával kiegészítve, ily módon sikerült 20% alá csökkentenünk a szóhiba-arányt, mely tudomásuk szerint a legalacsonyabb eddig publikált eredmény magyar nyelvű, nagyszótáras, folyamatos beszédfelismerés területén.

### 1 Bevezetés

A nemzetközi gyakorlatban a **folyamatos, nagyszótáras beszédfelismerő rendszerekben** (LVCSR – Large-Vocabulary Continuous Speech Recognition) tipikusan szóalapú nyelvi modellezést alkalmaznak. Azonban a morfológiailag gazdag nyelveknél – mint amilyen a magyar – e szóalapú megközelítés alkalmazása a jelentős szóalaki változatosság miatt megkérdőjelezhető. A klasszikus nyelvi modell az egyes N-gramok (szó N-esek) relatív gyakorisága alapján becsüli meg egy N-1 szóból álló előtörténet ("history") után álló szavak feltételes valószínűségét. Sok szóalak esetén kevés tanítóminta áll rendelkezésre egy kontextus becsléséhez, így a nyelvi modell döntése is kevésbé megalapozott különösen, ha akusztikailag nehezen megkülönböztethető szavak között kell választania.

Megoldást kínál a nyelvi modellek új, nyelvünkhöz jobban illeszkedő alapokra helyezése. Ehhez morfémaalapú nyelvi modelleket kell létrehozni a tanítószövegek szegmentálásával, majd a mintaillesztési folyamatot szavak sorozata ( $W$ ) helyett morféma sorozatán ( $M$ ) kell elvégeznünk. (1)

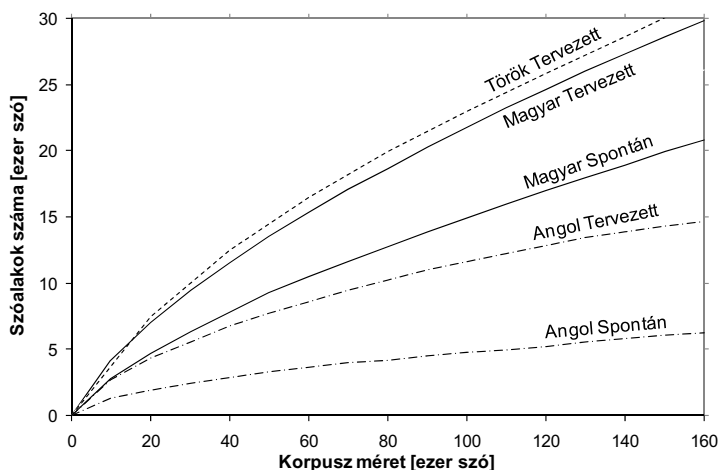
$$\hat{M} = \arg \max_M P(M) * P(O | M) \quad (1)$$

Ahol  $O$  a felismerendő beszédanyagból nyert jellemzővektorok sorozata.

A mintaillesztés morféma alapon is ugyanúgy elvégezhető, az egyetlen különbség, hogy a kimeneten is morfémasorozatot kapunk. Ezt újra szavakká összeilleszteni egyszerű feladat, ha előzetesen jelöltük a tanítószövegben a szóhatárokat. (2)

$$\hat{W} = f(\hat{M}) \quad (2)$$

A morfológiai gazdagságot jól jellemzi a felismerési feladat szótár-növekedési görbéje (1. ábra). Megfigyelhető, hogy az agglutináló nyelveknél, mint a magyar vagy a török mennyivel gyorsabb a szótár bővülés, és ennek üteme nem csak a nyelvtől, hanem a felismerési feladat jellegétől is erősen függ. Spontán beszédatadázison alkalmazva a morfémaalapú megközelítést korábbi munkánkban [1] csak kismértékű javulást értünk el a szó alapú rendszerhez képest. Ezzel szemben [2] jelentős hibaarány csökkenésről számol be török nyelvű, olvasott híryanagokon végzett kísérletek alapján. Morfémaalapú rendszerünkkel 20% körüli szóhiba-arányt értek el. Ennek háttérében az állhat, hogy a híryanag-felolvasás szóalakokban gazdagabb feladat, mint a spontán beszéd. Figyelembe véve, hogy a magyar és a török nyelv tervezett beszédatadázisokon nagyon hasonló szótárnövekedést mutat (1. ábra), okkal feltételezzük, hogy olvasott híryanagon, magyar nyelven is jobban teljesíthetnek a morfémaalapú felismerők.



1. ábra. Szótárméret növekedés a korpusz méretének függvényében  
(források - spontán magyar: [1]; török és angol eredmények: [3])

Cikkünkben öt különböző szegmentáló eljárást használó **morfémaalapú** felismerőt mutatunk be, melyek pontosságát egy standard, szó alapú felismerő pontosságával vetjük össze. Az említett rendszerek tanítása internetes híryanagok felhasználásával készült szöveges tanító-adatbázison történt. A teszteléshez egy kb. egy óra hosszúságú felolvasott híreket tartalmazó felvételt használtunk, mely egy országosan fogható televízió adásából került rögzítésre. A felismerési feladat részletes áttekintése után bemutatjuk hogyan nyertük a mintaillesztési folyamatban használt morfémákat, majd ismertetjük a tesztanyagon elért eredményeket. Végül összefoglalást adunk kísérleteink legfontosabb következményeiről.

## 2 A felismerési feladat

### 2.1 A szöveges tanító-adatbázis

#### A tanító-szöveg összegyűjtése

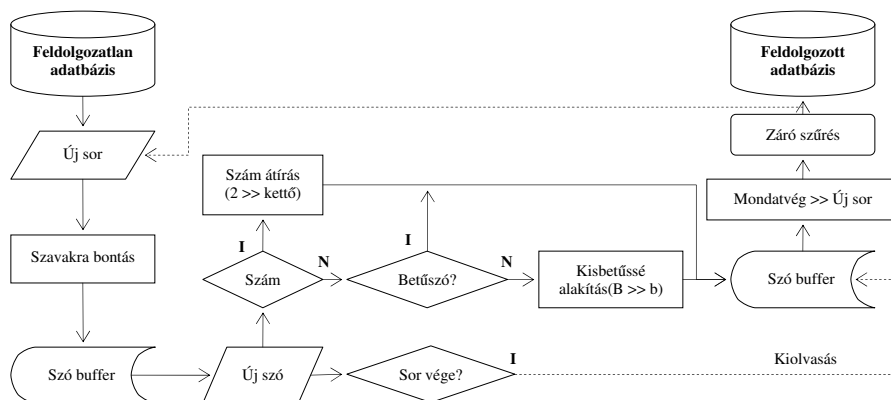
A cikkünkben ismertetett felismerő rendszer szöveges tanító-adatbázisa internetes gyűjtés eredménye, és egy országosan fogható televíziós csatorna portáljáról származik. Beszédfelismerők tanításnál a rendelkezésre álló beszédatadátbázist általában több részre osztják, aszerint hogy tanításra, tesztelésre vagy modell paraméterek hangolásra kívánják-e használni. Ebben az esetben általában a tanító beszédatatok leíratait használják fel a nyelvi modell tanításához. Ez azonban nagyon idő- és költségigényes, mivel kézi úton kell átírni a beszédjelet szöveggé.

Némi kompromisszum árán, de van hatékonyabb megoldás. Összegyűjthetők különböző forrásokból a felismerési feladathoz tematikában, szókincsben, struktúrában jól illeszkedő szöveges tanítóadatok. Az ilyen módon készített nyelvi modellek természetesen valamelyest kevésbé illeszkednek a konkrét beszédatadátbázishoz, azonban előnyük lehet, hogy gyorsabban, nagyobb méretben előállíthatóak, és kisebb célzottságuk miatt robosztusabb működést eredményezhetnek. Esetünkben az összegyűjtött tanítószövegek egy TV csatorna hírportálként is működő honlapjáról származnak. Az itt fellelhető belföldi híryanagok hat évre visszamenőleg kerültek összegyűjtésre, ami mintegy 54 ezer cikk feldolgozását jelenti. Ezekből állt össze a felismerő rendszer nyers tanító-adatbázisa.

#### A tanítószöveg előfeldolgozása

A beszédfelismerési feladat a szöveges tanítóadatok speciális előfeldolgozását követeli meg. Egy nyers internetes híryanag rengeteg olyan karaktersort tartalmazhat, melyeket nem kívánunk a nyelvi modellbe beépíteni, vagy a többi szótól eltérően szeretnénk kezelni. Az előbbire lehet példa a megjelenési dátum, forrásmegjelölés, esetlegesen bent maradt HTML tag, stb., mivel ezek csak feleslegesen rontják a nyelvi modell minőségét. Míg az utóbbira a betűszavak, vagy számok esete, melyekre a későbbiekben kitérek. Mindezek mellett az is elmondható, hogy azokat az írásjeleket is el kell távolítani a tanítószövegből, amelyeknek nincs egyértelmű megfelelője a hangképzésben, mint például a vessző, a kettőspont vagy a kötőjel. Az előfeldolgozás lépéseit szemlélteti a 2. ábra.

A nyers tanítószöveg feldolgozása soronként történik, **Perl** (Practical Extraction and Report Language) nyelven írt scriptek segítségével. Minden sor beolvasás után az egyes szavakon folytatjuk a feldolgozást. Erre azért van szükség, mert a számok és betűszavak az átlagos szavaktól eltérő eljárást igényelnek. Számok feldolgozásánál a nehézséget az jelenti, hogy nem rendelhető hozzájuk egyértelműen a kiejtett, fonetikus alakjuk. A probléma feloldásához azt a stratégiát választottuk, hogy átírjuk őket betűvel leírt alakjukra, így a felismerő ugyanolyan módon tanítható velük, mint a normál szavakkal. Ezt az átírást egy Perl szubrutin hajtja végre, tehát automatizáltan történik.



2. ábra. A tanító-adatbázis előfeldolgozása

A második megoldásra váró eset a betűszavaké. Ezekkel az a gond, hogy a magyar nyelvben egyszerű szabályokkal a kimondott alakjuk nehezen jósolható. Néha teljes értékű szóként ejtjük őket (pl.: APEH, KRESZ), néha viszont betűnként olvassuk ki (pl.: ÁNTSZ, DVD), nem is beszélve arról az esetről, ha idegen eredetű rövidítésről van szó (pl.: GDP, BBC). A számokkal ellentétben ezt egyszerű programozási eljárással nem lehet kezelni, így a fonetikai átíratukat a kiejtési modellben, kézzel adjuk meg. Ahhoz hogy ezt megtehessek, el kell kerülni a morféimákra bontásukat. Ezt legegyszerűbben úgy érhetjük el, ha meghagyjuk őket nagybetűs alakjukban, így formájukban elkülönülnek az átlagos szavaktól. Mindezek után minden egyéb szót kisbetűsítünk a tanítószövegben.

Hogy eljussunk a nyelvi modellezéshez alkalmas formához, már csak két lépést kell megtenni. Az elhangzó mondatokat egymástól függetlennek tekintjük a nyelvi modell szempontjából, ezért a mondatvégi írásjeleket „új sor” szimbólumra cseréljük. A második lépésben azt biztosítjuk, hogy a tanítószöveg végső alakjába ne kerülhesen semmilyen a felismerés folyamán nem értelmezhető karakter. Az ehhez alkalmazott szűrőfeltétel csak a magyar ábécé betűit engedi meg, minden más írásjelet töröl.

## 2.2 Akusztikai tanító-adatbázis

Az akusztikus modell tanításához összesen egy órányi átírt híryanag állt rendelkezésünkre. Ez önmagában kevés egy teljesen új akusztikus modell felépítéséhez, sőt mivel a teljes egy órát tesztelési célokra szerettük volna fenntartani a felügyelt adaptációról is le kellett mondanunk. Éppen ezért **beszélőfüggetlen** akusztikus modellként egy korábbi, a mostanitól független felismerési feladathoz illesztett modellt használtunk, mely eredetileg Magyar Referencia Beszédatadatbázison (MRBA) [4] lett tanítva.

Bár felügyelt adaptálásra nem volt lehetőségünk, felügyelet nélkül azonban végeztünk a rendelkezésre álló egy órás felvételt és a beszélőfüggetlen modell felhasználásával. Az így keletkezett akusztikus modellt használtuk **beszélőadaptált** kísérleteinkhez.

### 3 Morfémaszegmentálás

Mint a bevezetőben kitértünk rá a szóalapú nyelvi modellezés nehézségei főként nyelvünk szóalaki változatosságból erednek. A magyar nyelvben egyetlen szónak rengeteg képzett-ragozott formája létezik, így ugyanaz a szótő különböző kontextusban eltérő formákat vehet fel. Ennek következtében a tanítószövegben rendelkezésre álló információ elaprózódik, ami a szókapcsolatok pontatlan becslését eredményezi. E változatos morfológiával úgy küzdhetünk meg a legjobban, ha a szótári szavakat kisebb elemekre tudjuk bontani. Ha ezt a szegmentálást optimálisan hajtjuk végre, csökkenteni tudjuk a szótár méretét. Kisebb szótárméretnél a szótári elemek többször fordulnak elő, így több mintát szolgáltatnak a nyelvi modell tanításához, ami pedig végső soron hatékonyabb mintaillesztést tesz lehetővé.

A feladat elvégzéséhez azonban olyan módszerek bevezetése szükséges, amelyekkel a (2) képletben bevezetett  $f$  függvény inverze, a  $W$ -ről  $M$ -re képző  $f^{-1}$  optimálisan megvalósítható.

#### 3.1 Szabályalapú eljárások

A szegmentáló eljárások közül először a nyelvspecifikus szabályokon és szótáron alapuló módszereket mutatjuk be röviden. Kísérleteinkben a magyar nyelvű **Hunmorph** [5], általános célú annotáló rendszert használtuk, melynek részeként futásidejű morfológiai elemzésre az ún. Ocamorph program szolgál. Tudásforrásként a Morphdb.hu [6] adatbázist használja, mely minden eddigienél mélyebben megalapozott morfológiai leírását tartalmazza a magyar nyelvnek.

Fontos jellemzője a szabályalapú módszereknek, hogy az egyes szavak egymástól függetlenül kerülnek elemzésre, és az elemző általában több lehetséges szegmentálást is megad. A megfelelő kiválasztására valamiféle stratégiát kell alkalmazni. Ennek megfelelően két változata született az elemzésnek a szegmentáló program beállítása és a szegmentált alak kiválasztása szerint.

#### Hunmorph Compound-Guessing (HCG)

A morfológiai elemzés ezen változatában a szegmentálás minden módja megengedett. A futásidejű elemző felbonthatja az összetett szavakat (*--compounds* kapcsoló) sőt, ha nincs egy szóalagnak semmilyen érvényes elemzése, akkor egy az adatbázisban nem szereplő tagot is leválaszthat róla feltéve, hogy így elemezhető alakhoz jut (*--guess Fallback* kapcsoló).

#### Hunmorph Strict Fallback (HSF)

A szabályalapú morfológiai elemzés második változata a feldolgozás lépcsőzetességére épül. Az első lépcsőben a Hunmorph számára csupán az egyértelműen feldolgozható szavak elemzése engedélyezett. A második lépcsőben az így nem elemezhető szavakra megengedjük, hogy összetett szóként legyenek figyelembe véve. Végül az ezután is felbonthatatlan szóalakokról az adatbázisban nem rögzített elemek is leválasztásra kerülhetnek.

### 3.2 Statisztikai alapú eljárások

A Morfessor család tagjai Minimum Description Length (MDL) elven alapuló statisztikai eljárások, amelyeket finn kutatók fejlesztettek ki. A statisztikai alapú szegmentáló eljárások nagy előnye, hogy működésük nem igényel emberi felügyeletet. Nem használnak sem nyelvspecifikus lexikont, sem toldaléklistát, hanem a bemenetül kapott szótár statisztikai tulajdonságai alapján bontják a szavakat kisebb elemekre. Céljuk olyan optimális felbontást találni, mely tömören képes a korpuszt reprezentálni. Ennek következtében a szegmentálás eredményeként kapott morfémaszerű egységek (ún. **morfok**) nem feltétlenül rendelkeznek jelentéssel.

#### Morfessor Baseline (MB)

A Morfessor Baseline a finnek módszerének alapváltozata. Az optimális szegmentálás megkeresését a következőképpen önthetjük matematikai alakba. (3)

$$\arg \max_M P(M \mid \text{korpusz}) = \arg \max_M P(M) * P(\text{korpusz} \mid M) \quad (3)$$

Ahol  $M = \mu_1, \dots, \mu_m$  a korpusz egy lehetséges morf felbontását jelöli. A mintaillesztéshez hasonlóan itt is két paraméter értékének szorzatát kell optimalizálni ahhoz, hogy a korpusz legnagyobb valószínűségű szegmentálását megkapjuk: az adott  $M$  morfkészletnek, mint lexikonnak a valószínűségét ( $P(M)$ ) és  $M$  korpuszhoz való illeszkedési valószínűségét ( $P(\text{korpusz} \mid M)$ ). E két változó közelítésről bővebb leírás a kapcsolódó irodalomban található [7].

#### Morfessor Categories-MAP (MC-MAP)

A Morfessor család második tagja a baseline módszer finomításaként született statisztikai alapú szegmentáló eljárás. Legnagyobb újtásként az algoritmus megpróbálja kikövetkeztetni, hogy az egyes morfok prefixum, szótó vagy szuffixum szerepet töltenek-e be, és ezt a morf után helyezett címke (/PRE, /STM, /SUF) segítségével jelöli is a végeredményben. Bár ez tagadhatatlanul növeli a szótárméretet – ugyanis így ugyanaz a morf akár három lexikai elem szerepét is betöltheti – mégis megtérülhet ez a fajta megkülönböztetés a pontosabb nyelvi modellezésben. Részletekért lásd [8].

### 3.3 Hibrid eljárás

Mind a szabályalapú, mind a statisztikai alapú morféma szegmentálásnak lehetnek hátulütői. A Hunmorph-fal történő elemzések nem eredményeznek tömör szótárt, ami beszédfelismerési feladatnál nem előnyös, hiszen minél több szótári elem között kell a dekódolás során különbséget tenni, annál nagyobb a hiba lehetősége is. Ezzel szemben a statisztikai eljárások hatásfokát a nem elégséges mennyiségű tanítóadat ronthatja le. Éppen ezért felmerült az igény a két szemlélet egyesítésére.

A **Combined Hunmorph-Morfessor (CHM)** eljárás a HCG szabály alapú módszer szegmentálásán alapszik. Lényegében a MB algoritmus átalakítása oly módon, hogy a szegmentálás valószínűségi becslését a HCG által szolgáltatott felbontási

alternatívákon végzi el. Így tehát csak olyan morfémák keletkezhetnek, melyek a szabály alapú eljárással jöttek létre, viszont a statisztikai módszer biztosítja, hogy arra az alternatívára essen a döntés, mely globálisan tömör morfémakészletet eredményez. Az ehhez szükséges algoritmust [9] részletezi.

## 4 A beszédfelismerő hálózatok kiértékelése

### 4.1 Beszédfelismerési paraméterek és beállítások

Minden felismerő rendszer ugyanazon a kb. egy órás olvasott híryanagon lett kiértékelve. A 16 kHz-en mintavételezett és 16 biten kvantált felvételek lényegkiemeléséhez dinamikus Delta és Delta-Delta értékkel kiegészített mel-frekvenciás kepsztrális komponenseket (**MFCC** – Mel-frequency cepstral coefficients) és ún. vak csatorna-kiegyenlítést alkalmaztunk. Akusztikus modellként balról-jobbra struktúrájú, három-állapotú rejtett Markov-modelleket használtunk, állapotonként 7 Gauss függvényből álló sűrűségfüggvényekkel.

A szóalapú nyelvi modell tanítása a normál, előfeldolgozott tanítószövegen, míg a morfémaalapú felismerők esetén, ugyanezen szöveg szegmentált változatain történt. Minden felismerő Kneser-Ney simított [10], trigram nyelvi modellen alapszik, mely az SRI-LM [11] nyelvi modellező toolkit segítségével lett előállítva. Entrópia alapú modell-metszést [12] csak a szó alapú hálózatban alkalmaztuk, memória takarékosági szempontból (1. táblázat).

1. táblázat: A tanító- és tesztadatbázis adatai

	Szavak száma [ezer szó]	Szóalakok száma [ezer szó]	OOV arány [%]
Tanító-	5600	285	–
Tesztadatbázis	7.6	3.5	3.6

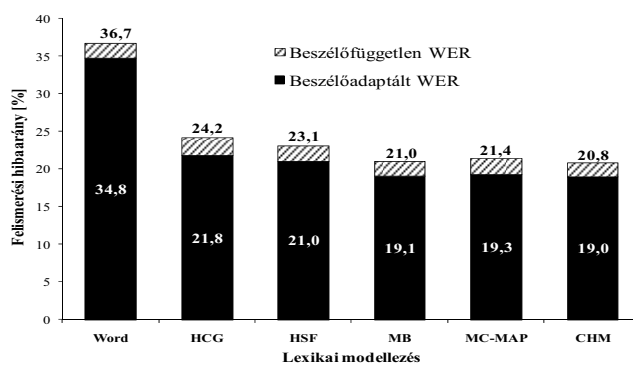
A felismerő hálózatok építését az Mtool WFST (Weighted Finite State Transducer) keretrendszer programjaival végeztük, míg a tesztelés során alkalmazott egyutas mintaillesztési feladathoz a VOXserver [13] nevű dekódert használtuk. A dekódolási folyamat számításigénye az ún. Real Time Factor (RTF) tekintetében az egyes feladatok között kiegyenlítésre került.

### 4.2 Felismerési eredmények

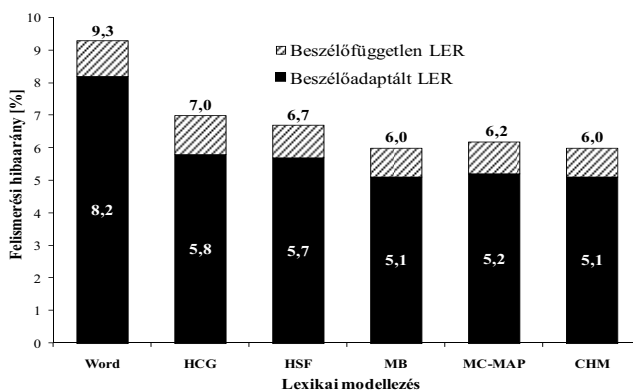
A felismerő rendszerek teljesítményének értékeléséhez szóhiba-arányt (**WER** – Word Error Rate), illetve betűhiba-arányt (**LER** – Letter Error Rate) számoltunk. Ezen kívül feltüntettük a szó alapú kiindulási rendszerhez képest elért relatív WER csökkenéseket beszélőfüggetlen és beszélőadaptált esetben is (2. táblázat, 3. ábra).

2. táblázat: Beszédfelismerési eredmények

Technika	Szótár mérete	Beszélőfüggetlen eredmények			Beszélőadaptált eredmények		
		WER [%]	$-\Delta\text{WER}_{\text{rel}}$ [%]	LER [%]	WER [%]	$-\Delta\text{WER}_{\text{rel}}$ [%]	LER [%]
Word	285 e	36.7	–	9.3	34.8	–	8.2
HCG	50 e	24.2	37	7.0	21.8	37	5.8
HSF	63 e	23.1	34	6.7	21.0	40	5.7
MB	31 e	21.0	43	<b>6.0</b>	19.1	45	<b>5.1</b>
MC-MAP	45 e	21.4	42	6.2	19.3	45	5.2
CHM	80 e	<b>20.8</b>	43	<b>6.0</b>	<b>19.0</b>	45	<b>5.1</b>



3.1 ábra. Beszélőfüggetlen és beszélőfüggő szóhiba-arányok



3.2 ábra. Beszélőfüggetlen és beszélőfüggő betűhiba-arányok



### 4.3 Értékelés

Az eredmények ismeretében elmondható, hogy a morfémaalapú felismerők szignifikánsan (Conf.=95%) jobban teljesítettek a felismerési feladaton, mint a szó alapú rendszer. Az újonnan bevezetett eljárásokkal elért átlag 40%-os relatív felismerési hiba csökkenés figyelemre méltó eredmény, és egyben felhívja a figyelmet arra, hogy nyelvünk jobban modellezhető morféma alapon egy szóalakokban gazdag feladat esetén. A bevezetett módszerek közül a legnagyobb felismerési pontosságot a hibrid szegmentáló eljárással (CHM) sikerült elérni, de fontos megjegyezni, hogy ettől szignifikánsan a statisztikai megközelítésekkel (MB, MC-MAP) nyert eredmények sem térnek el. A statisztikai és hibrid rendszerekhez képest viszont szignifikánsan magasabb felismerési hibát kapunk a szabályalapú szegmentálások (HCG, HSF) alkalmazásakor. Érdekes megfigyelni a beszélőadaptáció hatását is, miszerint jellemzően tovább növeli a szó és morfémaalapú felismerő közötti pontosság különbséget [14].

Vizsgáljuk meg, mitől pontosabbak a morfémaalapú rendszerek. Morféma alapon OOV szavak is felismerhetők. Míg a szó alapú felismerő csak a szótárában található szavakat képes helyesen felismerni, addig a morfémaalapú rendszerekben a lexikai elemek elvben tetszőlegesen összekapcsolódhatnak, így a tanítószövegben nem szereplő szavak is előállhatnak a mintaillesztés folyamán. Figyelembe véve azonban, hogy a tesztanyag mindössze 3.6%-a OOV szó, ez önmagában nem adhat választ az ennél jóval magasabb abszolút hiba csökkenésre. A döntő tényező valójában, a szó alapon fennálló adatéltelenség kezelése, aminek köszönhetően a szótáron belüli szavak felismerésekor is jóval kevesebb helyettesítési és törlődési hiba keletkezik.

## 5 Összefoglalás

Cikkünkben öt különböző szegmentáló eljárással készített morfémaalapú gép felismerő teljesítményét hasonlítottuk egy standard, szó alapú rendszeréhez tervezett beszédű, híryanag felolvasásos feladaton. Minden újonnan bevezetett módszerrel szignifikáns, átlag 40%-os relatív hibaarány csökkenést sikerült elérnünk, mely a morfológiailag gazdag feladat pontosabb nyelvi modellezésére vezethető vissza. Különösen jól teljesítettek a statisztikai alapú szegmentáló technikák, ezen belül is legkiemelkedőbben egy hibrid eljárás, mely nyelvspecifikus tudást is felhasznált. Felügyelet nélküli adaptációs technológia segítségével a szóhiba-arányt 20% alá tudtuk szorítani, mely tudomásunk szerint egyedülállóan alacsony magyar nyelvű, LVCSR feladaton.

Korábbi munkáink [1],[15] is bizonyították, hogy más nyelvekhez hasonlóan magyar nyelven is eredményesen alkalmazható a morfémaalapú nyelvi modellezés, azonban ilyen mértékű javulás egyetlen korábbi feladat esetén sem volt mérhető. Ennek oka az lehet, hogy a szóalaki változatosság és relatív hibaarány csökkenés erős kapcsolatban áll egymással. Minél gazdagabb morfémaalkotókban a felismerési feladat, annál nagyobb szükség van olyan lexikai modellezés használatára, mely a szavaknál alacsonyabb szintű nyelvi elemeket is figyelembe veszi.

## Köszönetnyilvánítás

Ezúton szeretnénk köszönetet mondani az AITIA International Zrt.-nek és a THINKTech Kutatási Központ Nonprofit Kft.-nek a rendelkezésünkre bocsátott eszközökért és adatokért. Kutatásunkat részben az OM-00102-2007-es projekt támogatta.

## Hivatkozások

1. Mihajlik, P., Tüske, Z., Tarján, B., Németh, B., Fegyő, T.: Improved Recognition of Spontaneous Hungarian Speech – Morphological and Acoustic Modeling Techniques for a Less Resourced Task. *IEEE Transactions on Speech and Audio Processing* (megjelenés alatt)
2. Arısoy, E., Can, D., Parlak, S., Sak, H., Saraçlar, M.: Turkish Broadcast News Transcription and Retrieval. *IEEE Transactions on Audio Speech and Language Processing*, vol. 17, no. 5 (2009) 874-883
3. Creutz, M. et. al.: Morph-Based Speech Recognition and Modeling Out-of-Vocabulary Words Across Languages. *ACM Transactions on Speech and Language Processing*, vol. 5, Issue 1, Article no. 3 (2007)
4. Vicsi K., Kocsor A., Teleki Cs., Tóth L.: Beszédatadabázis irodai számítógép-felhasználói környezetben. In: II. Magyar Számítógépes Nyelvészeti Konferencia, Szeged (2004)
5. Trón, V., Németh, L., Halácsy, P., Kornai, A., Gyepesi, Gy., Varga, D.: Hunmorph: open source word analysis. In: *Proc. ACL 2005 Software Workshop* (2005) 77-85.
6. Trón V., Halácsy P., Rebrus P., Rung A., Simon E., Vajda P.: Morphdb.hu: magyar morfológiai nyelvtan és szótári adatbázis. In: III. Magyar Számítógépes Nyelvészeti Konferencia, Szeged (2005)
7. Creutz, M., Lagus, K.: Unsupervised Morpheme Segmentation and Morphology Induction from Text Corpora Using Morfessor 1.0. In: *Comp. and Inf. Sci.*, report A81, HUT (2005)
8. Creutz, M., Lagus, K.: Inducing the Morphological Lexicon of a Natural Language from Unannotated Text. In: *Proc. of AKRR'05*, Espoo, Finland, 15-17 June (2005)
9. Németh B., Mihajlik P., Tikk D., Trón V.: Statisztikai és szabály alapú morfológiai elemzők kombinációja beszédfelismerő alkalmazáshoz. In: V. Magyar Számítógépes Nyelvészeti Konferencia, Szeged, (2007)
10. Chen, S. F., Goodman, J.: An Empirical Study of Smoothing Techniques for Language Modeling, Technical Report TR-10-98, Computer Science Group, Harvard University (1998)
11. Stolcke, A.: SRILM – an extensible language modeling toolkit. In: *Proc. Intl. Conf. on Spoken Language Processing*, Denver (2002) 901–904
12. Stolcke, A.: Entropy-based Pruning of Backoff Language Models. In: *Proc. DARPA Broadcast News Transcription and Understanding Workshop* (1998) 270-274
13. Fegyő, T., Mihajlik, P., Szarvas, M., Tatai, P., Tatai, G.: VOXenter - Intelligent voice enabled call center for Hungarian. In: *EUROSPEECH-2003* (2003) 1905-1908
14. Tüske Z., Mihajlik P., Fegyő T., Trón V.: Spontán, nagyszótár, folyamatos beszéd gépi felismerési pontosságának növelése beszélőadaptációval a MALACH projektben. In: V. Magyar Számítógépes Nyelvészeti Konferencia, Szeged (2007)
15. Tarján B.: Large-Vocabulary Continuous Speech Recognition in Hungarian. In: *Végzős Konferencia 2009, Budapest, 2009. május 20.* (2009)